National Human Genome Research Institute

Summary of Boston Strategic Planning Town Hall – Establishing a 2020 Vision for Genomics

Wednesday, July 17th, 2019

Overview

On Wednesday, July 17, 2019, the National Human Genome Research Institute (NHGRI) held an in-person strategic planning town hall at Massachusetts General Hospital in Boston, Massachusetts. The goal of this public meeting was to solicit input for NHGRI's strategic planning process. As of summer 2019, NHGRI was past the half-way point of this strategic planning process and pivoted from gathering open-ended input to seeking feedback on major themes that have emerged from the process. Towards that end, the town hall focused on emerging themes in the areas of genomic variation identification, association and function in human health and disease. About 120 individuals from academia and industry attended this event.

Opening Remarks and Welcome

The session opened with remarks from Michael Talkowski, Ph.D., of Massachusetts General Hospital (MGH). Dr. Talkowski pointed out that many cross-institution National Institutes of Health (NIH) programs, including some large NHGRI programs, have participation by researchers in Boston. So, it is particularly appropriate that attendees of this town hall contribute to NHGRI's strategic planning process.

NHGRI Director Eric Green, M.D., Ph.D., presented an overview of NHGRI's current strategic planning process, including a history of the institute's previous planning activities. The goal is to publish the new NHGRI strategic plan in October 2020, which would coincide with the 30th anniversary of the launch of the Human Genome Project. Dr. Green highlighted that human genomics research is now supported far more by other NIH Institutes than by NHGRI. By 2020, NHGRI will only be funding about 10% of the human genomics research at the NIH. Dr. Green asked: where should NHGRI focus in the future? What is the forefront of genomics? The NHGRI strategic plan will guide NHGRI priorities and define what NHGRI means by the forefront of genomics.

Following introductory remarks, the town hall featured three sessions: 1) basic genomics and genomic technologies, 2) genomics of human health and disease, and 3) a trainee-only session. There was an extended question-and-answer session with Dr. Green, as well as other NHGRI representatives, including:

- Carolyn Hutter, Ph.D., Director of NHGRI's Division of Genome Sciences.
- Elise Feingold, Ph.D., Scientific Advisor for Strategic Implementation in NHGRI's Division of Genome Science.
- Adam Felsenfeld, Ph.D., Program Director in NHGRI's Division of Genome Sciences.
- Elaine Ostrander, Ph.D., NHGRI, Chief of NHGRI's Cancer Genetics and Comparative Genomics Branch.
- Bill Pavan, Ph.D., Chief of NHGRI's Genetic Disease Research Branch.

Basic Genomics and Genomic Technology

Elise Feingold, Ph.D., NHGRI, introduced the audience to the concept of the five internal NHGRI strategic planning focus groups, two of which were the focal points of the town hall: Basic genomics and Genomic Technologies and Genomics of Human Health and Disease. Dr. Feingold described previous events that provided input, including town halls, internal focus groups, expert advisors' groups, the National Advisory Council for Human Genome Research, and strategic planning meetings, specifically the "From Genome to Phenotype: Genomic Variation Identification, Association and Function in Human Health" workshop, held in January 2019 in Bethesda, Maryland.

Dr. Feingold described the first theme ("Enable facile, routine generation of whole-genome sequences & transcriptomes and characterization of epigenomes") as foundational work that is critical for understanding genome function. The second theme ("Understand and interpret whole-genome sequences, transcriptomes and epigenomes") is critical, because there is still much that is not understood about how identical (or near-identical) DNA sequence gives rise to the specialized functions of specific cells, tissues, and organs and how these functions are influenced by genomic variation and environment. The third theme ("Establish the role(s) of all genes and regulatory elements in pathways, networks and phenotypes") is very important, as we know that genes and regulatory elements do not work in isolation. The fourth theme ("Use evolutionary and comparative genomic data to markedly advance understanding of genome function") focuses on a renewed vision for comparative genomics. The fifth theme ("Enable facile, routine generation and use of synthetic nucleic acids in genomics research studies") focuses exclusively on an emerging area for technology development. The sixth theme ("Understand and leverage population structure and admixture to facilitate human genetics studies") is a significant challenge relevant to all aspects of genomics. In the context of basic genomics and genomic technologies, individual admixture, which is very common in the U.S. and worldwide, is a confounding factor in most genomics studies ranging from basic research to disease analysis.

Discussion

Town hall attendees liked the framework of NHGRI at the forefront of genomics and viewed the institute as having a continued role in supporting cutting-edge genomics research.

Participants emphasized the need for standardized formats for sequence and variation data from large cohorts. Participants suggested that NHGRI should take the lead in developing and implementing such formats. NHGRI representatives were asked how NHGRI envisions the shift from "short read genomes" to "long read genomes." Renewal of the NHGRI genome reference program will lead some of this, but it is important to determine what quality level is needed for clinical applications. The genomics research community have done a lot of production on the current model and will need to balance cost versus specific needs in the future. Stakeholders also recommended that it be made clear that all the themes contain a genetic variation component.

Participants also noted that determining the role of all regulatory elements in pathways and phenotypes is very important. We do not yet have robust, well-settled methods for thinking about all regulatory elements in the true network context in which they exist. Establishing the role of all genes and regulatory elements in pathways, networks and phenotypes at scale is a massive amount of work. NHGRI hopes to be able to make a quantum leap in this area. What is the right way to model this concept that engenders excitement, while still being feasible and attainable? Participants suggested that taking on a model organism, or just a pathway, or developing standards would be a way to move forward. What are the technologies, resources and methods that are needed?

NHGRI may want to separately capture physiological data and make it relatable to the many types of "-omics" studies. Study participants need to be part of the feedback loops in order to correctly capture phenotype. The biggest barrier to epigenomics studies has always been getting access to tissue. What is the extent to which this is left up to the hospitals and biobanks, versus NHGRI? We need well characterized tissues that many researchers can access. Biobanks must be engaged.

Participants questioned whether there is a way to interface with local scientists and groups to address the diversity issue? Different communities are sensitive to inclusion in research. NHGRI is thinking about those concerns across focus groups and partnering with others at NIH to work on best practices with different populations.

Attendees pointed out that an important nuance in the language that NHGRI uses is about gene regulation and not non-genic sequence. NHGRI uses "elements" to be more inclusive than just genes, but participants argued for NHGRI to consider adding to the strategic plan more specific language outside the linear genome, such as the non-gene regulatory genome and structural genome.

Genomics of Human Health and Disease

Adam Felsenfeld, Ph.D., NHGRI, described each of the themes from Genomics of Human Health and Disease. The scaling work under the first theme ("Establish the functional consequences of genomic variants affecting human health and disease") is NHGRI's rational next step for Genomics of Human Health and Disease. It includes both variant function and gene function, covers a range of scales and applications, and a variety of contexts. The second theme ("Determine the genomic architecture of

human diseases and traits") includes activities that are already underway, but not complete; more examples are needed to understand generalities. Genomic architecture is a framework for understanding the variant-to-phenotype relationship. The third theme ("Develop the methods and analyses to support use of non-sequence genomic data for characterizing human health and disease") includes transcriptomes, modified bases, chromatin structure, proteomics, metabolomics and environmental data. The fourth theme ("Transform how we assemble sample sets for genomic studies of human disease") depends on high-quality samples that are broadly consented and which cover a range of human phenotypes. The fifth theme ("Commit to systematic inclusion of appropriate ancestral diversity into all large-scale genomic studies and analyses") is essential, since under-representation of diverse human populations in genomic research and resources will hinder genomic science and ultimately result in disparities.

Discussion

Participants discussed how it is essential to increase sample sizes in relevant studies. Investigators have difficulties convincing study sections of this issue. The most effective thing that we can do early is to make sure that data are easily related across phenotypes. This involves data standards and places to go to standardize data; scientists need to see the value of an added dataset. NHGRI can pitch to other NIH Institutes the value of increased samples sizes, but it is important to be able to point to the results and other lines of evidence. NHGRI is small in terms of its funding, but NHGRI recognizes the need to take a leadership role, when appropriate, and educate others.

One attendee pointed out that the *All of Us* precision medicine program will have a large and diverse genomic data set. *All of Us* fits in under "transforming how we assemble sample sets." How will NHGRI capitalize on this and other efforts to move the genomics community agenda forward? How will NHGRI work with *All of Us* in this larger forum and integrate the views of the involved parties?

Standards development for clinical sequencing will need to be addressed. In the past, standards for data sharing were developed in research centers as sequencing itself developed. Participants challenged NHGRI to consider how the institute can continue a standard model of data sharing and analysis as sequencing moves away from being generated by NHGRI centers.

The assigned study section for an application influences the scoring of that application and ultimately whether it gets funded. Do we need new study sections to address emerging challenges? This is an issue for fields that are shifting. NHGRI sees "variation to function to disease" as a critical area; it is anticipated that NHGRI will be developing FOAs that can influence where applications are reviewed.

Due to the size of NHGRI's budget, NHGRI cannot sequence everything. However, there is a great long-standing relationship between the sequencing centers executing large projects and NHGRI. NHGRI could still be at the forefront of how the large projects are managed (organizing committees, data sharing, etc.). There are benefits from data sharing and analysis using NHGRI's way of doing things.

Participants gave feedback on how to frame the themes, arguing that they should relate more to biological problems to be tackled. Taking common, complex or rare disease as an example, one can see how genetic variation teaches us about the biology of disease. If questions were framed around that problem, investigators could figure out how to take the complex or rare disease and determine key genes, cell types, cellular pathways and networks. The questions then become: how could NHGRI build resources in ways to connect diseases to biological functions? What key resources and computational methods are needed?

Other Topics of Discussion

Participants raised social aspects related to people having more knowledge of genetic information and knowing what diseases they might get. While not explicitly presented at this town hall, the NHGRI planning efforts include a focus group that is addressing society, education and engagement issues. NHGRI continues to recognize ethical, legal and social research as important areas and is engaging with healthcare systems, industry and other stakeholders on these topics.

Regarding genomic medicine, participants discussed the need for research in economically diverse settings, coordination with international partners, reviewers with experience in clinical settings, and scaling technologies for use in clinical settings.

Discussions on data science issues included: models to compute directly where the data is stored, a single model to reduce redundancies, central organization, definition of sustainable data, and the importance of active curation of information.

Trainee Session

A session dedicated to trainees was held immediately after the town hall. The trainee session was moderated by Susan Slaugenhaupt, Ph.D., MGH. Carolyn Hutter, Ph.D., NHGRI, outlined background for the strategic planning process.

Trainees asked how much of the “Forefront of Genomics” is meant to be pathbreaking. NHGRI intends to support risky research that others will not, as well as more foundational research and resources that will enable others to study human genomics. The group discussed where rare disorders fit. NHGRI has not explicitly excluded them. NHGRI considers a broad range of disorders.

NHGRI has done a lot of work bringing together functional datasets. These could also link back to the DNA sequence and perhaps other levels of data as well. NHGRI and NIH recognize the high priority of such a resource. OMIM, GeneReviews and other sources of knowledge are key resources, especially for trainees. What is the NHGRI view on keeping these resources available and funded for trainees to learn genetics? NHGRI funds OMIM and is working to establish funding from other sources for

genomics/informatics resources in general. Resources that are developed must be available to train the next generation of scientists.

NHGRI provided advice on several career development topics, including:

- trainee and fellowship award mechanisms.
- the new NHGRI Genomic Innovator award that targets genomics researchers who have trained in a big collaborative environment and have not gotten their first R01 yet.
- grantsmanship and the value of resources, including examples of how to write good grants.
- how Program Directors can be a great source of information for navigating the NIH grant process.
- balancing risk in applications, particularly in early career stages.
- best practices for establishing collaborations.